

## 「統計分析」についての補足解説資料

京都大学大学院教授 藤井聡

本補足資料では、まず重回帰分析の概要を簡潔に解説しつつ、政府分科会が行った重回帰分析の結果、および、その公表資料を解説する。その上で、筆者が証言において取り上げた「検出力」「単位根過程」について解説する。なお、本資料は筆者が纏めたものであるが、本資料が一般的な解説から乖離していない事を示す主旨にて筆者とは別の者がまとめた資料、**参考資料 1 (重回帰分析)**、**2 (検出力)**、**3 (単位根過程)** もあわせて提出する。

## ■重回帰分析、について

① (東京都が引用した) **政府分科会**が行った分析は、「**重回帰分析**」と呼ばれる分析です。

②一言で、重回帰分析とは何かを説明すると次の様になります。

- a) まず、**目的変数  $y$**  を考える。
- b) 一方で、その**目的変数  $y$**  が**複数の説明変数  $x_1, x_2, x_3, \dots$**  によって、「**説明される**」と**仮定**する。
- c) そして、重回帰分析とは、以上の仮定の下で
- c1) 「**想定した目的変数  $y$  が、説明変数  $x_1, x_2, x_3, \dots$  によって本当に説明されているのか、否か?**」、そして、
- c2) もしも説明されているのなら、「**どういう風に説明されるのか?**」を、**統計学的に分析する方法論**です。いわば、「**要因分析**」を行うことが目的です。
- d) なお、その**要因分析結果**に基づいて、「**予測**」を行うこともできます。

③ここで、重回帰分析における、**上記 b) の「説明変数で目的変数を説明する」**とは、例えば、次の様な事を意味しています。

$$\text{実効再生産数} = 1.2 + 0.12 \times (\text{21 時の人出}) + 0.15 \times (\text{気温}) + \varepsilon$$

これはつまり、

- 「**実効再生産数の値は、**
- ・ (21 時の人出) に 0.12 掛け合わせたものと、
  - ・ (気温) に 0.15 掛をけ合わせたものに、
  - ・ 1.2 を足し合わせたものである」

という事を意味しています (なお、ここで、 $\varepsilon$  は誤差項と呼ばれるものなのですが、ここではその説明を割愛します。簡単に説明すると、平均ゼロで一定の確率分布をもって変動する変数、と言うことになります)。

なお、この例では**実効再生産数**が「**目的変数**」、**21 時の人出**や**気温**は「**説明変数**」です。

そして、「0.12」や「0.15」が「**係数**」と呼ばれます (1.2 は切片あるいは定数項と呼ばれます)

「**重回帰分析**」では**まず最初に**、この 0.12 や 0.15 といった**係数** (および切片) を**データに基づいて計算 (推計)** します。

④ここで、この数式（一般に回帰式と言われます）が正しければ、「21 時の人出が増えれば、実効再生産数が増える」という結論が、統計学的に導かれる事になります。だから、もしも、上記の回帰式が正しければ、感染症対策として 21 時の人出を減らすことが正しい、という事になります。

ところが一方で、この回帰式が正しくなければ、感染症対策として 21 時の人出を減らすことが正しいとは言えない、という事になります。

⑤では、「この回帰式の正しさ」をどうやって判定するか、というと、そこで用いられるのが、「統計的検定」(statistical test) という手法です。

⑥「統計的検定」とは「近代的な裁判」と同じ構造を有しているもので、一般に次の様なものです。

a) まず最初に、「検証すべき仮説」を想定する（例：21 時の人出が増えれば、感染が拡大する[いわば、有罪]）

b) その「検証すべき仮説」を「対立仮説」として、それとは逆の仮説（つまり、それが否定されれば対立仮説を採択せざるを得ない仮説）を「帰無仮説」とする（例：21 時の人出は、感染に影響しない[いわば、無罪]）。

c) その仮説を検証するためのデータを集める（例：実効再生産数と 21 時の人出の時系列データ）

d) 「仮に、帰無仮説が正しいと想定した場合に、そのデータが得られる確度」を求める。そしてこの確度を有意確率  $p$  値と呼ぶ。（例：「21 時の人出は、感染に影響しない（いわば無罪）」ことを前提とした場合に、当該データが得られる可能性を  $p$  値とする）

（なお、 $p$  値を計算するために、重回帰分析の係数の検定の場合には、 $t$  値というものを算定する。これは、 $p$  値計算のために算定される統計量である。係数検定には、平均値の検定の折りに採用される  $t$  検定が行われるために、そのための統計量として  $t$  値が算定されることになっている。 $t$  値とは、得られたデータに基づいて想定される当該係数の推定値についての「確率分布」の期待値と 0 との間の距離を標準化した値であり、そして  $p$  値とは、当該確率分布において、その分布における確率変数とその期待値から、 $t$  値以上離れる確率を意味する）

e) 一方で、「 $p$  値が、この確率を下回ったら、帰無仮説が正しいという想定が間違っているのではないか？」と想定する確率を、予め決めておく。一般にその確率を「有意水準」あるいは「危険度」と呼称する。一般的に、この確率は高い場合でも 5%、より安全な判断を目指す場合は 1% とする。

f) その上で、以上の d) で求めた  $p$  値と、e) で想定した有意水準を比較する。そして、

- ・  $p$  値が有意水準（5%あるいは 1%）を下回れば、帰無仮説を棄却する。  
そして、「対立仮説が統計学的に有意だ（統計学的に効果がある）」と判断する。  
（例：統計学的に言って 21 時の人出は実効再生産数に減らす効果を有意に持つ  
Or 21 時の人出は、実効再生産数に対して統計的有意な効果を持つ）
- ・  $p$  値が有意水準（5%あるいは 1%）以上ならば、帰無仮説を棄却しない。  
そして「対立仮説が統計学的に有意でない（統計学的に意味はない）」と判断する。  
（例：統計学的に言って、21 時の人出は実効再生産数に影響を与えているとは言えない  
Or 21 時の人出は、実効再生産数に対して統計的有意な効果を持たない）

なお、以上の f) の帰無仮説を棄却するか否かの判断が一般的に「統計学的検定」と言われる。

⑥重回帰分析では、各説明変数について推計した各説明変数の係数について、以上の統計学的検定を、一つ一つ行っていきます。

⑦以上に加えて、重回帰分析では、『説明変数全体』で、「目的変数の散らばり」をどれだけ説明できたのか？』を計算します。その結果は、 $R^2$ （アール二乗値、あるいは、アールスクエア：決定係数）と呼ばれるもので、0～1 の値を取り、より大きい程、説明変数全体で、目的変数の散らばりをよく説明できる、という事になります。なお、 $R^2$  は一般に、説明変数が多いほど大きくなる傾向がありますが、その傾向を割り引いた決定係数は、「自由度調整済み決定係数  $adj.R^2$ 」と呼ばれます。

（なお、「 $R$ 」は「重相関係数」と呼ばれ、説明変数から推計される目的変数と、目的変数の実測値との間の相関係数を意味します。また、 $R$ 、 $R^2$  と同様の重回帰分析を全体として評価する  $F$  値というものもあります）。

この  $R^2$  や  $adj.R^2$  が 1 に近ければ、用いた説明変数群で目的変数を「予測」することに活用する事も可能となりますが、これが 0 に近ければ「予測」することが難しいという事になります。しかし、仮に  $R^2$  や  $adj.R^2$  が 0 に近くても、それはその重回帰分析が行っている検定結果が無効であることを全く意味しません。ただ、説明変数群で目的変数の散らばりを殆ど説明できない、という統計学的事実を意味するに過ぎません。むしろそれは、説明変数の効果の無さを積極的に意味するものにはなりません。

## ■政府分科会が行った「実効再生産数の増加」についての重回帰分析結果

①以上が、重回帰分析の概要の説明ですが、まとめて言いますと、重回帰分析を行った場合、その結果を読み解くには、最低限、以下の数値を報告することが必要です。

- ・重回帰式がどうなったか？（すなわち、**各説明変数の係数**）（および切片）
- ・その重回帰式の「係数」は統計的に有意か否か？（**各説明変数の p 値**）（および t 値）

②東京都が被告準備書面(3)にて、時短命令の実証的根拠としてあげた政府分科会資料の重回帰分析結果における上記情報は公開されていなかったため、内閣官房に直接問い合わせたところ、「実効再生産数の増加」について、次の様な結果を、政府分科会が得ていることが分かりました。

### 「新規陽性者数の変化率（実効再生産数に対応）」を目的変数とする重回帰分析結果

	係数（偏回帰係数）	t 値	p 値
8 時の人出	-0.098	-0.72	0.47
気温	-0.42	-3.00	0.013
21 時の人出	0.25	1.55	0.13

サンプルサイズ=38  $R^2=0.37$

これを見ると、

- ・8 時の人出、21 時の人出は、統計的に有意な影響を持っていない事が分かります。  
（理由は p 値がそれぞれ 0.47、0.13 でいずれも 0.01 のみならず 0.05 よりも大きいからです）
- ・気温は統計的に有意な影響を持っていることが分かります。  
（理由は、p 値が 0.013 で、0.05 よりも小さいからです）

これはつまり、8 時にしろ 21 時にしろ「人出」が増えたからと言って、感染者数の変化率が増えるということは、統計学的有意には全く言えない事を意味しています（あるいは、人出が感染者数の変

化率に影響を及ぼすという仮説は、統計学的に棄却されている、したがって、人出が感染者数の変化率に与える影響は統計学的に意味があるものではない事を意味しています。

なお、この数値を内閣官房が持っていたということは、以上の統計学的事実、すなわち、「21時の人出は感染者数の変化率に対して影響を持つという仮説が、統計学的に棄却されているという事実」を、政府分科会が知っていた事を意味します。

③なお、政府分科会が、この重回帰分析について「公表」している情報は、下記のものです。

〈新規陽性者数の増加局面における寄与率〉

- 新規陽性者数の増加局面においては、忘年会の寄与率が大きかったと考えられる。このことはツイッター分析と一致。
- また、人出の中でも、特に21時の人出の寄与率が大きかったと考えられる。

増加局面（2020年12月5日～2021年1月11日）における寄与率試算

		目的変数	説明変数1	説明変数2	説明変数3	説明変数4
東京	変数パターン1	新規陽性者数の変化率	8時の人出	飲み会	カラオケ	忘年会
		---	9%	-19%	29%	44%
	変数パターン2	新規陽性者数の変化率	8時の人出	気温	21時の人出	
		---	-13%	-54%	33%	
大阪	変数パターン1	新規陽性者数の変化率	8時の人出	飲み会	カラオケ	忘年会
		---	14%	-4%	30%	53%
	変数パターン2	新規陽性者数の変化率	8時の人出	気温	21時の人出	
		---	2%	-54%	44%	

(注) 各種データを元に内閣官房で試算

(注1) 増加局面（2020年12月5日～2021年1月11日）について、東京都及び大阪府の新規陽性者数の変化率に対する14日間の人出（8時、15時、21時）、気温・温度、  
 (注2) 飲み会のツイート数（全国データ）の寄与率を重回帰分析により試算した。なお、疫学的な要因は考慮していない。  
 (注3) 上の表の中では、用意した複数の説明変数の組み合わせのうち、有益と考えられるその分析結果の一部を掲載している。

(注3) Twitterデータ（Twitter, NTT データより提供）は、東京大学豊田正史教授による統計データを使用。

ご覧の様に、t 値、p 値、係数のいずれもが公表されていません。したがって、この公表情報では、「21時の人出」が効果を持つのかどうかを解釈不能です。

ただし、(注2) では、「用意した複数の説明変数の組み合わせのうち、有益と考えられる分析結果の一部を掲載している」と記載されており、有意か否かは不明だが、政府分科会が、これら変数を説明変数として用いることを「有益」だと考えていることが分かります。

また、公表されているのは、「寄与率」と呼ばれるものだけでした。この寄与率という概念は、一般的な重回帰分析では用いない概念です。ついては、内閣官房に問い合わせたところ、「係数の絶対値の総和に占める、その変数の係数の比率」とのことでした。

■重回帰分析の係数についての検定における、「検出力」について

①ところで、重回帰分析の係数の「検定」の精度は100%ではありません。「本来なら効果がある」（＝有罪な）のに、「有意で無い」と判定（無罪判決）するという「過誤」もあり得ます。

※例えば、この度の裁判で、東京都側は、我々が行った分析で時短、まん防、緊急事態宣言には、有意な感染抑止効果は「ない」という判定（＝無罪判決）は、ただ単にその効果が統計的に有意な水準に達しなかっただけで、本来は、そういう効果があった（＝有罪であった）可能性があるのではないかと指摘しています。つまり、東京都は、我々の判定に「過誤」があるではないかと指摘しているわけです。

②そういうリスク（＝有罪なのに無罪判決を出してしまうリスク）がどれくらいあるのか無いのかを

考えるために算定されるのが「検出力」です。

③**検出力**とは、対象とする検定毎（＝判決毎）に算定されるもので、その定義は、『その検定で、本当は効果（＝本当は有罪）がある場合に、その検定で、しっかりと「有意」（＝有罪）だと判定できる確率』です。したがって、最低は0で最高は1の値をとり、1に近いほどに、その検定において「有意で無い」という判断（＝無罪判決）の確度が高い、ということになります。

④ただし、(特定のサンプルサイズと有意水準を想定した) 個々の検定において、その**検出力を想定するには、「検出すべき効果量」を定義する必要があります**（これはつまり、「どういう水準の罪科があるか？」を定義することに対応します）。

⑤例えば、当方が提出した資料には、以下の表が掲載されています。この表には、

検出力 (0.2) と 検出力 (0.1)

という二つの「列」がありますが、これは

「0.2 (あるいは 0.1) という効果量が実際に存在する場合に、その効果量を正しく検出できる確率」を意味しています。例えば、「赤点線の○」で囲んだ所に 1.0 と書かれていますが、これは、

“「無規制から時短等へダミー」という変数が、

実際に 0.2 の効果量を持っていたとした場合、

この回帰分析で、その効果を「棄却」せずに「有意」だと判定する確率が 1.0 である “

という事を意味しています (なおこの 1.0 は四捨五入値であり実際には 1 未満です)。言い換えるなら、「真の効果量が 0.2 の場合、その効果を正しく検出できる確率は (ほぼ) 100%である」、さらに逆に言うなら、「間違っ

て棄却してしまう過誤が生ずる確率は (ほぼ) 0%である」ということを意味しているのです (ちなみに、0.1 の場合の検出力も「1.0」となっていますから、0.1 でも正しく検出できる確率がほぼ 100%だという事になります)。

説明変数	非標準化係数	標準誤差	t	p	検出力(0.2)	検出力(0.1)
(Intercept)	0.00	0.00	-0.27	0.79		
無規制から時短等へダミー	0.03	0.02	1.16	0.25	1.00	1.00
時短等からまん延防止へダミー	-0.02	0.02	-0.87	0.39	1.00	1.00
時短等から緊急事態へダミー	-0.01	0.02	-0.44	0.66	1.00	1.00
まん延防止から緊急事態へダミー	0.01	0.02	0.53	0.60	1.00	1.00
緊急事態からまん延防止へダミー	0.01	0.02	0.43	0.67	1.00	1.00
緊急事態から時短等へダミー	0.00	0.02	0.05	0.96	1.00	1.00
[従属変数: 実効再生産数の前日差]					調整済みR <sup>2</sup>	-0.01

そして、この表の検出力は全て 1.0 となっていますから、全ての変数について 0.1、0.2 の効果があればそれを正しく検出できる確率がほぼ 100%であることを意味しています。つまり (効果量として 0.1, 0.2 を想定すれば)、東京都が裁判で指摘した「この回帰分析で間違っ

て棄却してしまう」という確率は「ほぼ 0%」であり、東京都の指摘はほぼあり得ないという事を示しています。いずれにしても、これだけ高い検出力の下で分析し、有意な効果が見られなかった、という我々の上記表の分析結果は「本当に効果がない可能性が極めて高い」という事を意味しています。

### ■単位根検定について

①今回、政府分科会が行ったのは、「実効再生産数と 21 時の人出の、連日の値の推移のデータ」です。

こうしたデータは、一般に「**時系列データ**」と言います。

- ②そして、時系列データの中には、「近い過去の値」と強い関係を持っているものがあります。例えば、毎年の日本の人口は、昨年的人口に基づいています。これは、  
**「前期の値に『変化量』を加えて、今期の値が決まる」**  
という時系列データで、こうしたものは特に「**単位根過程**」と呼ばれます（その典型的なものは、「ランダム・ウォーク」データと呼ばれるもので、「千鳥足の酔っ払いの位置データ」がその一例となっています）。
- ③こうした単位根過程の時系列データを、「前期の値の影響」を無視して分析すると、何の意味も無いのに、「統計的に有意な相関」が検出されてしまうことがあります。こうした無意味な相関は一般に「**見かけの相関**」と呼ばれるものです。したがって、**単位根過程の時系列データを、そのまま統計分析することは、統計学的な時系列分析においては「御法度」となっています。**
- ④したがって、そのため**時系列データを分析する際には、それが「単位根過程か否か」を必ず確認します。そして、単位根過程では「ない」と判断された場合に限り、それ以後の分析に進みます。**（一般に、そういう過程は“**定常過程**”と呼ばれます）。
- ⑤なお、その時系列データが単位根過程か否かを判断する方法が、「**単位根検定**」です（これは、その時系列データが、単位根過程で「ある」という帰無仮説を措定し、それがデータによって棄却されるか否かを統計学的に判断する検定です）。つまり、**単位根検定で「単位根過程でないと判断できない」という結果になった場合には、そのデータをそのまま統計分析することは統計学では御法度とされるのです**（検定しても見かけの相関しか得られない恐れが大きいためです）。
- ⑥ただし、単位根過程の時系列データでも、前期とのデータとの「差」を求めれば、その差の時系列データは、単位根過程では「ない」という可能性が出てきます。したがって、通常の時系列分析では、**単位根である可能性が示された時系列データについては、前期との「差分」をとり、それに対して上記の④、⑤、⑥の検定・分析を繰り返します**（ただし、通常一回目の差分（例外的に二回目の差分）をとって単位根過程が棄却できなければ、そのデータは適切な分析ができないデータであると見なして、それ以上の分析を停止することが一般的です）。
- ⑦なお、政府分科会が「増加局面」と「減少局面」のそれぞれについて回帰分析を行った実効再生産数データについて単位根検定を行った結果、「増加局面」においても「減少局面」においても、単位根過程である可能性が示されました。したがって、**政府分科会は一般的な統計学においては御法度とされている「単位根過程の時系列データに対する分析」を行ってしまっていた訳です。**これでは、**彼等が見いだした統計的傾向が仮にあったとしても、それが「見かけの相関」に過ぎない可能性が濃密に危惧されるのです。**事実、差分をとって分析したところ、彼等の結論とは逆に「21時の入出は実効再生産数の下落に全く貢献しない」ことが明確に示されています（詳細は、当方が証人として出廷した折りに提出した「東京都の時短命令と緊急事態宣言の感染抑止効果について」を御参照下さい）。